

**THE USE OF INTELLIGENCE AND ACHIEVEMENT TESTS  
IN THE PRIMARY SCHOOL**

**GLENNIS E. SCHULTZ**







### Manuscript Theses

Unpublished theses submitted for the Master's Degree are deposited in the Florida Southern College Library and are available for inspection. Use of any thesis is limited by the rights of the author. Bibliographical references may be noted, but passages may not be copied unless the author has given permission. Credit must be given in subsequent written or published work.

A library which borrows this thesis for use by its clientele is expected to make sure that the borrower is aware of the above restrictions.

Florida Southern College Library



THE USE OF INTELLIGENCE AND ACHIEVEMENT TESTS  
IN THE PRIMARY SCHOOL

GLENNIS E. SCHULTZ

Submitted in partial fulfillment of the  
requirements for the degree of Master  
of Arts in the Graduate School of  
Florida Southern College

August, 1949





APPROVAL

---

Professor Edward Lee Flemming, Advisor

Donald A. Thompson


Reader

---

Reader

August 10, 1949

Date Submitted to the Chairman  
of the Graduate Committee



Digitized by the Internet Archive  
in 2011 with funding from  
LYRASIS Members and Sloan Foundation

## ACKNOWLEDGEMENTS

The writer wishes to express her gratitude to all who have given advice and assistance in the preparation of this theses: to Professor E.L. Flemming, advisor; to Professor D.A. Thompson and Dr. C.L. Murray, readers; and to her parents Mr. and Mrs. C.D. Schultz.



# TABLE OF CONTENTS

CHAPTER	PAGE
I. SETTING THE PROBLEM . . . . .	1
The problem . . . . .	1
General statement . . . . .	1
Specific problems . . . . .	3
Definition of terms . . . . .	3
Intelligence tests . . . . .	3
Achievement tests . . . . .	4
Primary school . . . . .	4
Delimitations . . . . .	4
Basic assumptions . . . . .	5
Basic hypothesis . . . . .	5
Procedure in collecting data . . . . .	5
Procedure in treating data . . . . .	6
The need for the study . . . . .	6
II. REVIEW OF THE LITERATURE . . . . .	10
Early history . . . . .	10
Intelligence tests . . . . .	13
Achievement tests . . . . .	18
First tests in Pinellas County . . . . .	20
III. THE ELEMENTARY TESTING PROGRAM . . . . .	23
Selection of the tests . . . . .	25
Frequency of testing . . . . .	29
IV. THE TESTS . . . . .	32



CHAPTER	PAGE
The California Test of Mental Maturity . . .	32
The Metropolitan Achievement Tests . . . . .	36
V. ADMINISTERING AND SCORING THE TESTS . . . . .	40
Administering . . . . .	40
Scoring . . . . .	43
VI. THE VALUE OF TESTING IN THE PRIMARY SCHOOL . .	48
VII. SUMMARY AND CONCLUSIONS . . . . .	55
BIBLIOGRAPHY . . . . .	59





## CHAPTER I

### SETTING THE PROBLEM

#### The Problem

##### General Statement

The purpose of this study is to investigate and to evaluate the use of intelligence and achievement tests in Primary Schools of St. Petersburg, Florida.

An alert teacher is always, either consciously or unconsciously, studying pupil reactions and consequently developing within herself either favorable or unfavorable attitudes toward the children. Since informal and casual study of behavior reactions may lead to inaccurate and prejudiced opinions, it is necessary that the teacher be helped to develop and use objective and controlled techniques of diagnosis. Observation which is definitely planned, unprejudiced and accurately recorded, rather than casual and disorganized, will favorably minimize the "halo" effect.

Certain of these techniques, such as the Stanford-Binet Test of Intelligence, the Bernreuter Personality Inventory, the Downey Will-Temperament Test, and physiological measurements of the emotions, because of their complexity, elaborateness, or delicacy of measuring techniques, should be administered only by the trained expert or in the psychological laboratory. Other measuring techniques may be

---

<sup>1</sup>Lester D. Crow and Alice Crow, Mental Hygiene in School and Home Life. pp. 273-274.



administered safely by the intelligent, careful, and trained teacher, and their results interpreted.<sup>2</sup>

Standardized tests used in the Primary School yield a diagnostic profile that provides an analysis of the abilities of each pupil. They also have comparable norms for the various abilities. The teacher can utilize the results of the tests in the improvement of her instructional program in many ways: in planning a remedial program to meet the needs of the pupil; in grouping pupils according to their needs for remedial instruction; in discovering the difficulties of each child in the skills which are essential for success in pupil activities; and in providing a basis for studying the adjustment problems of the child.

A general intelligence test furnishes a means of determining a pupil's ability to profit by training and instruction.

The measures of intelligence, aptitude, and readiness, if used with other types of data, furnish a scientific and practical basis for selection and placement of learning material. They assist not only in determining the difficulty and complexity of material that the individual is capable of learning, but supply a valid basis for long-range planning with reference to educational and vocational goals.<sup>3</sup>

This project will be a study of the basic testing

---

<sup>2</sup>Ibid. pp. 273-274.

<sup>3</sup>Harl R. Douglass, The High School Curriculum. p. 56.



program of the St. Petersburg area with the thought that the over-all picture might prove useful in improving, intensifying, reconstructing, or revising the present system. A survey of the system should prove useful in a logical inspection of the program and should aid in creating uniformity of administration and evaluation.

### Specific Problems

It is the purpose of this study to determine the value of a Central Testing Control of the Primary Schools of St. Petersburg, Florida through investigation of (1) the intelligence and achievement tests in use in the Primary Schools of this city; (2) the frequency of administration of these tests; (3) reasons for the selection of the particular tests; (4) administration of the tests; and (5) interpretation of the results of the tests.

### Definition of Terms

Intelligence Tests. Tests designed to measure the degree of one's general alertness without emphasis upon any specific ability. These tests may be used by the teacher to:<sup>4</sup>

1. Provide a diagnostic evaluation of the mental abilities of each child
2. Interpret learning difficulties in the light of pupil abilities

---

<sup>4</sup>J. Murray Lee and Dorriss May Lee, The Child and His Curriculum. pp. 624-625.





3. Identify outstanding pupils of either high or low ability who need special educational procedures
4. Analyze individual problem cases, such as failing pupils, disciplinary problems, and personality difficulties
5. Understand the general mental level of the class as a guide in the selection of instructional materials<sup>5</sup>

Achievement Tests. Tests of a survey or diagnostic type which measure the results of teaching and learning.

- (1) They are so formed as to test only those things that can be more or less objectively tested.
- (2) Only the definite achievements in the subject studied are tested, as far as this is possible.
- (3) Each test is arranged so that the answers are always given in the same way, a definite and precise length of time is given for the test, the explanations and directions to be given are always the same, and definite, detailed directions for scoring the test are given.<sup>6</sup>

Primary School. Includes the first through the sixth grades.

#### Delimitations

Because the testing program for the Primary Schools in the St. Petersburg area is determined not by the individual school, but by a central testing bureau, no attempt will be made to analyze a specific school situation, but rather the

---

<sup>5</sup>Ibid., pp. 624-625.

<sup>6</sup>Arthur J. Jones, Principles of Guidance. p. 142.





subject will be treated from a general, or policy, standpoint. As indicated in the general statement, the scope of this investigation will be limited to the St. Petersburg, Florida area.

### Basic Assumptions

Because of a significant difference in the native capacity, abilities, and interests of the individual student, personal consideration by the teacher of the various differences is a definite factor in the classroom situation.

If a testing program is to be of value, the test administrator should have a background of definite and specific training in the testing field. Proper training will insure proper administration and evaluation of intelligence and achievement tests.

### Basic Hypothesis

It is the intent of the writer to show that a Central Testing Bureau is of value in the Primary School, and that such a control will lead to a more valid selection of tests, improved administration of these tests, and to a more careful interpretation of the test results.

### Procedure in Collecting Data

The facts and information contained in this research are based on the current situation--that is the conditions existing in the 1947 - 1948 school year. The research to be contained herein is concerned solely with the testing program



in the St. Petersburg, Florida area.

There is a multitude of sources for material and it would be impossible to list specifically the sources available. Publications, pamphlets, and assorted printed material are available through commercial firms, through State and Federal bureaus, public and private agencies and professional associations and societies.

Since a great part of the information contained in the thesis was obtained locally, some data was obtained through personal interview with local school faculties and from the members of the staff of the Central Testing Bureau located in Clearwater, Florida.

#### Procedure in Treating Data

In treating the above data the current Primary School testing situation was examined in all its phases and compared, where possible, to the theoretically ideal situation. Defects and deficiencies are noted where observed and reference made to the proper methods as determined by our currently recognized specialists in the field. Where possible, practical suggestions for improvement are offered.

#### The Need for the Study

Recently the writer was called upon to aid in the administration of a battery of tests in one of St. Petersburg's Primary Schools. Realizing how insufficiently prepared she was to administer properly these tests at that



time, the writer decided to investigate the testing situation in the elementary schools of St. Petersburg. Because there is little published literature concerning the worth of a Central Testing Bureau, or of the use of the administration of a thorough battery of tests in the primary grades, the writer chose to use the material and data of her investigation as the subject of her thesis.

In all the primary grades, as well as the more advanced grades, there are children who, because of low intelligence, are unable to advance as rapidly as other children of the same age. It is necessary to determine which children in the primary school have insufficient intelligence to profit from the usual instruction. It would be doing the children an injustice to segregate them on the basis of the teachers' judgments; many teachers would have prejudices or would be subjected to various influences from parents and others, and therefore could not be relied upon to make objective judgments. The best method, therefore, to determine children needing remedial help would be a series of tests administered regularly over a period of years.

Because there is a definite need for testing in the primary school, it is necessary to determine which tests would be best suited so that the childrens' intelligence and achievement might be measured without the influence of teachers and parents. There is also a need to discover a child's mental development so that he might be compared with





a normal level, and thus determine by how many years he is retarded or advanced.

Through this study the writer intends to show how the results of the intelligence and achievement tests administered to children in the primary grades may be used to aid the teacher, the principal, the child, and the parents. In choosing a particular test it will be shown that several important factors must be taken into consideration, such as the ease of administration, of scoring, of interpretation and application, the cost, the proper mechanical makeup, and the requirements the test must meet to be of value.

A study of the use of the intelligence tests will also be of value in rectifying some of the faulty conceptions which people hold about Intelligence.<sup>7</sup> There are many of these improper beliefs, and it is necessary to recognize the true nature of ability, and to evaluate childrens' abilities and then to plan an education suitable to these abilities.

One of the false conceptions commonly held concerning intelligence is that the degree of intelligence can be estimated from the general appearance of an individual's facial characteristics. Head size, certain types of mannerisms and specialized behavior, such as knitting of the forehead, are believed by some people to indicate intelligence.

---

<sup>7</sup>Harry J. Baker, Introduction to Exceptional Children, pp. 224-225.





Others believe that speed of movement, or rapid talking are indications of mental ability. Only through the use of carefully administered and evaluated tests can the true degree of intelligence be estimated.



## CHAPTER II

### REVIEW OF THE LITERATURE

#### Early History

For centuries men have been puzzled over differences between individuals in their ability to learn, to reason, and to get along in their various stations of life. The mentally backward were met, in the medieval ages, with scorn, ridicule, and often very barbaric forms of punishment, or "treatment". In contrast, the mentally gifted have been honored and praised for their accomplishments. The mystery of human abilities has been the cause of much philosophical study in the past, and of much psychological study at present.

Among the foremost contributions toward educational advancement in the twentieth century has been the development of a system of measurement of individual differences. Though a beginning in the accurate scientific measurement of these differences was not undertaken until the end of the nineteenth century, the differences have been noted throughout educational history.

Plato (ca. 428 B.C. - ca. 348 B.C.)<sup>1</sup>, in his Republic, sorted out and prepared children to become philosophers, warriors, or artisans according to their talents. He was in-

---

<sup>1</sup>Robert Ulich, History of Educational Thought.p. 1.



terested in training only the elite, and ignored the common man and practical and vocational interests. Like Aristotle, he believed that "The development of an individual is largely determined by hereditary factors--not only physically, but also with respect to his psychic qualities."<sup>2</sup> Quintilian (ca. A.D. 35 - 95)<sup>3</sup> admitted that some boys surpassed others in ability, and that dull, unteachable ones were rare; and Vittorino da Feltre (1378 - 1446)<sup>4</sup> stated that no one should expect education for careers which were beyond their capacity.

John Locke (1632 - 1704)<sup>5</sup> advanced the theory of "tabula rosa". This theory reduced the mind to a blank at birth, or to equality. His French followers accepted this theory and maintained that the minds were not only blank, but possessed of the same sensibility, differences due to different environments and educations. It was, of course, quite easy to notice these differences, but very difficult to assay them. Quite often this was accomplished through recitation or examination, which resulted in the separation of the bright students from the dull in both intellectual aptitude and scholastic achievement. During the eighteenth century noting differences by

---

<sup>2</sup>Ibid. p. 37.

<sup>3</sup>Ibid. pp. 51-60.

<sup>4</sup>Ibid. p. 110.

<sup>5</sup>John S. Brubacher. A History of the Problems of Education. p. 160.





these methods was prominent, and there is evidence of a means of segregation by means of terms 'best, second best, inferior, and worse'. During the nineteenth century this became a numerical system, which later led to decimals.

A demand led to a more exact system of measuring individual differences, and one of the first to experiment was Herbart (1176 - 1841)<sup>6</sup>, who maintained that the dynamic quality of life expressed in idea, feeling, and will was supposed to be the result of the mutual attraction or repulsion of reals as they effected each other. Because he lacked objective and experimental evidence for his conclusions, his attempt to measure psychological data in mathematical terms was unsuccessful. In 1833 Galton<sup>7</sup> outlined a method for studying free association by quantitative methods and made an outstanding contribution in statistical analysis when he suggested a graphical method of representing correlation.

In 1879, William Wundt<sup>8</sup>, father of experimental psychology, established the first experimental laboratory in Liepzig. His only interest in measurement was confined to reaction times, with little notice of the problem of individual differences.

---

<sup>6</sup>Ibid., p. 161.

<sup>7</sup>Ibid., p. 161.

<sup>8</sup>C.C. Ross, Measurement in Today's Schools. p. 36.





## Intelligence Tests

Psychological differences were studied by Cattell (1860 - )<sup>9</sup> who began the modern period of psychological testing. In Wundt's laboratory, in 1879, he discovered that different individuals had different rates of reaction time. The term "mental tests", which became a trade-mark for the measurement movement, was suggested by Cattell in 1890. He confined his tests, like Galton, to the simpler mental processes such as sensory discrimination, where individual differences are least rather than to the mental processes, where they are greatest; his tests were of the first attempts to measure intelligence.

Separate tests, which were later used by investigators in constructing their scales for measuring general intelligence were introduced by other German psychologists such as Kraepelin, Ebbinghouse, and Meumann.<sup>10</sup> "The possibility of applying statistics to education, however, was not perceived until Thorndike and his followers familiarized American educators with such concepts as averages, means, sigma scores, probable errors, correlations, and the like."<sup>11</sup> It was Thorndike who was responsible for educators becoming measurement conscious, and who offered the first university course in

---

<sup>9</sup>Ibid. pp. 41-42.

<sup>10</sup>Ibid. p. 36.

<sup>11</sup>Brubacher, op. cit., p. 36.



measurement, which was largely an application of statistics to education.

In 1897, when J. M. Rice<sup>12</sup> startled the educators of this country, at a meeting of the National Education Association, with the proposal that the results of teaching spelling could be measured by a spelling test, he was met with sarcasm and ridicule. However, opposition rapidly disappeared and a means of determining the kind and amount of work which a school system was doing was introduced in the form of a series of carefully devised "Standardized Tests". "The underlying purpose of the new movement began the creation of standardized scales for measuring school work, and for comparing the accomplishments of different schools and groups of school children."<sup>13</sup>

The momentum and scope of measurement increased rapidly. In France, interest in psychology and in testing had always been great, and there the problem of measuring intelligence was approached from the standpoint of the classification and treatment of the mentally defective.<sup>14</sup> Outstanding in the history of testing is the Frenchman Alfred Binet.

---

<sup>12</sup>Walter S. Monroe, Educational Tests and Measurements.

<sup>13</sup>Ibid. p. vii.

<sup>14</sup>Ross, op. cit., p. 39.



Binet (1857 - 1911) studied in several fields before becoming established in that of Psychology and Testing.<sup>15</sup> He first studied law, then medicine, and later worked in a biological laboratory. In an effort to devise a satisfactory method of measuring intelligence he tried experiments in physiognomy, graphology, palmistry, and in head measurements. The French government commissioned Binet and his assistant Simon to work out a test so that authorities might know better what children were of such low intelligence that they should be committed to an institution.<sup>16</sup> In 1905 they produced the first scale for the measurement of intelligence, which has served as the pattern for all subsequent tests and scales. This scale was made up of a series of selected questions to test common sense and judgment within a field of experience common to all children of a given culture. By giving the test to numbers of children they were able to establish norms for children of various chronological ages.

The degree of retardation was determined by subtracting the intellectual age from the chronological. Only later was it seen by another investigator that a more constant relation between the two factors could be gained by dividing the intellectual by the chronological age. This gave the intelligence quotient, or I.Q., so widely used to measure normal and accelerated children as well as retarded ones.<sup>17</sup>

---

<sup>15</sup>Ibid. p. 39.

<sup>16</sup>Brubacher, op. cit., p. 162.

<sup>17</sup>Ibid. p. 162.





In America, Goddard<sup>18</sup> was the first psychologist to recognize the value of Binet and Simon's scale. He translated it, and with minor adaptations, tried it out at Vineland. By 1910 Huey<sup>19</sup> had also made a translation; and in 1912 Kuhlman<sup>20</sup> published his first revision, with the difference that he extended it downward to the age of three months, rather than the three years which was Binet's lower limit.

The first thorough revision of the Binet Scale was published in 1916 by Terman<sup>21</sup> of Stanford University. Adapted to and standardized for use with normal as well as subnormal American children, it became known as the Stanford Revision or Stanford-Binet, and appeared with a complete manual "The Measurement of Intelligence".

Despite the criticisms that the Stanford Revision was standardized entirely on school children, which might handicap those of poor academic background, and that its norms were based on the children of only one state, California, the Stanford-Binet has been used for over twenty years, and is regarded as the best individual intelligence test. In 1937 Terman and Merrill prepared and published revisions of the Stanford Revision.

---

<sup>18</sup>Ross, op. cit., p. 42.

<sup>19</sup>Ibid., p. 42.

<sup>20</sup>Ibid., p. 42.

<sup>21</sup>Ibid., p. 42.





The Binet scales and later revisions had two drawbacks. First, they were chiefly verbal; Second, they were designed to be administered to individuals. In order to fill these deficiencies, steps were taken by various psychologists. Ebbinghaus<sup>22</sup> devised "completion tests" which have come to be accepted as measures of general intelligence; and Pintner and Paterson<sup>23</sup> assembled a series of fifteen tests of manipulation or performance to be used with deaf children. Known as the Pintner-Paterson Performance Scale, these tests appeared in 1917.

At this time America found herself involved in World War I.<sup>24</sup> The need arose to select commissioned and non-commissioned officers, and to meet this need the American Psychological Association offered its services to the War Department. Existing tests were not qualified to test illiterates or foreign speaking recruits, either as a group or individually. By utilizing the yet unpublished tests of Otis, a system of group tests was devised, and came to be known as the Army Alpha.<sup>25</sup>

Although the Army Alpha was a group test, it still

---

<sup>22</sup>C.C. Ross, Measurement in Today's Schools. p. 44.

<sup>23</sup>Ibid. p. 44.

<sup>24</sup>Ibid. pp. 44-46.

<sup>25</sup>Henry E. Garrett, Great Experiments in Psychology, Chapter 2.



emphasized the verbal factor. In order to meet the needs of testing illiterates and foreign speaking persons, the Army Beta, a completion test, was then devised. Later, group tests of the performance type were designed specifically for use with young children just entering school. Specific tests of intelligence, which emphasized capacity in a restricted area or school subject were developed. Ross states that "There is evidence that the development of the future is likely to be along the line of tests for specific intelligence, or aptitude in a restricted area, rather than tests of general intelligence, which aim to cover the whole range of human capacity at one shot."<sup>26</sup>

#### Achievement Tests

The testing of achievement is a much newer field than that of testing intelligence. One of the first to take an interest in a form of achievement test was an English Reverend, George Fisher.<sup>27</sup> In 1864 he arranged a "Scale Book" made up of various specimens arranged in order of merit. However, his work produced no lasting results; probably because progress in the scientific study of education was not yet advanced enough to offer statistical methods with which to carry on investigations.

---

<sup>26</sup>Ross, op. cit., p. 48.

<sup>27</sup>Ibid., p. 48



Shortly after Fisher's "Scale-Book" appeared, Rice introduced his spelling tests<sup>28</sup>; he was followed by Thorndike, who, through his experiments, became known as the father of the testing movement.

Besides his pioneer work on intelligence and his publications on statistical methods in education, Thorndike was responsible for most of the early standard tests and scales for measuring achievement. The first scale was the Thorndike Handwriting scale, published in 1910; and the first test was the Stone Arithmetic Test, published in 1908.

Scales and tests appeared rapidly within the next few years. Lending impetus to the movement was the school-survey movement, the appearance of books and periodicals related to the field, and the fact that educators were just beginning to realize how bad existing measurements were. One of the most striking studies was that of Starch and Elliott<sup>29</sup>, who in giving the same geometry paper to a number of pupils by different teachers discovered that school marks are highly subjective, given more as a function of the personality of the instructor than as a performance record of the student.

In 1918 Monroe<sup>30</sup>, in an investigation, discovered evi-

---

<sup>28</sup>Supra, p. 11.

<sup>29</sup>Ross, op. cit., pp. 50-53.

<sup>30</sup>Ibid., p. 54.





dence of a growing conviction that emphasis should be put upon quality of work rather than on quantity. This discovery hastened advancement of new developments in achievement tests.

The earlier forms of achievement tests were of the survey type. These gave a general measure of the pupil's achievement but did not give the information necessary for remedial work. By 1927 achievement tests were developed of a specific type. Diagnostic tests, which gave information on a pupil's strong and weak points; and practice tests, which emphasized drill, were introduced. Men contributing to this phase of the testing movement were Pintner, Monroe, and Buckingham.<sup>31</sup> The Stanford Achievement Test appeared in 1922, and with successive revisions, is still in use today.

#### First Tests in Pinellas County

The first county-wide testing program was in October and November, 1925. It consisted of a battery of ten educational tests and the Otis Group Intelligence Test: Advanced Examination. "More recent data could have been selected, but at that time there had been very few adjustments made, hence repeaters were more numerous."<sup>32</sup>

In regards to testing in the elementary school of today, J. Murray Lee states that:

---

<sup>31</sup>Ibid., p. 54.

<sup>32</sup>G.B. Fuguitt, A Study of Pupil Achievement in Pinellas County Public Schools, Clearwater, Florida. p. 44.





The determination of what is happening to boys and girls as a result of school experiences has come to be known by the term evaluation. The main concept guiding evaluation programs is that evaluation should be in terms of the extent to which pupils have attained the objectives of elementary education . . .

A great difference exists between the usual testing program of a school and the program which would be developed if the school put this thesis into practice. The testing program of most schools deals with the achievement of children in reading, arithmetic, and language usage. It is usually limited to certain abilities and knowledges. Obviously this is not sufficient to obtain an evaluation of the objectives which these elementary schools will claim they are trying to meet.

When schools are faced with this discrepancy, there is a wide variety of reasons which are given to account for it. 'Measurements of the objectives are not available,' 'It takes too much time,' . . . yet there are many more measurements available than are ever used . . . Achievement in subjects is thus revealed to be the real and only objective of many schools.

Where teachers in a school have actually tried to formulate their objectives and attempt to measure them, the results have been excellent. New measuring devices have been utilized; new material and teaching procedures have been introduced into the school."<sup>33</sup>

A testing program is essential in the elementary schools of today. Although men have thought about individual differences for centuries, it was not until the end of the nineteenth century that a beginning in the scientific measurement of these differences took place. Today, testing is still a growing field and it has already become one of

---

<sup>33</sup>J. Murray Lee, The Child and His Curriculum, pp. 595-596.



one of the most important means of gaining information about the individual, of predicting the possible outcome of his various efforts, and of classifying him in school into slow-learning or quick-learning groups.



## CHAPTER III

### THE ELEMENTARY TESTING PROGRAM

The word 'program' is indicative of order, system, and planning. A good testing program should be the result of careful thought and planning. Spence<sup>1</sup> has suggested that "a good testing program should be supplementary not duplicative, usable not confusing, economical not burdensome, comprehensive not sporadic, suggestive not dogmatic, progressive not static."

The testing program used in the primary grades of St. Petersburg, Florida, is the same as that carried on throughout Pinellas County. This county-wide system of testing is directed by a Central Testing Bureau which has its main office located in the School Administration Building, Haven Street, in Clearwater, Florida. The branch office for St. Petersburg is located in the Mirror Lake Junior High School and the supervisor divides her time each week between the two offices. Among her duties are the following:

1. To define the problem and clearly indicate the specific goals of the testing program.
2. To select the appropriate tests, preferably in cooperation with principals or committees of principals and teachers.<sup>2</sup>

---

<sup>1</sup>Ralph B. Spence, "A Comprehensive Testing Program for Elementary Schools." pp. 279-284.

<sup>2</sup>Gertrude H. Hildreth, Manual for Interpreting. p. 12.







3. To order the tests and related materials well in advance of the date on which they are to be used.
4. To formulate a schedule for the testing program.
5. To provide a set of directions outlining the administrative details of the program.
6. To distribute or arrange for the distribution of tests and test materials.
7. To train the principals or teachers in administering and scoring the tests.<sup>3</sup>

The testing program is geared to the facilities available within the school for administering, scoring, and interpreting the results. The teachers are given a bulletin prepared by the supervisor which contains explanations and dates of administering the tests. By supplying the teachers with an adequate knowledge of tests and testing techniques standardized tests assume an essential place in the work of the teacher.

The testing program is a long-range program. This makes it possible to have a varied program without leaving gaps or involving needless duplication. It also assures the taking care of present needs of the pupil as well as the needs of his future years. Also, the program is designed to supplement the ordinary informal tests and examinations given by the classroom teacher. Ross sets up the following eight steps, or stages, which the ordinary program should

---

<sup>3</sup>Ibid., p. 12.



consist of:

1. Determining the purpose of the program.
2. Selecting the appropriate test or tests.
3. Administering the tests.
4. Scoring the tests.
5. Analyzing and interpreting the scores.
6. Applying the results.
7. Retesting to determine the success of the program.
8. Making suitable records and reports.<sup>4</sup>

### Selection of the Tests

The selection of the tests, as mentioned before, is a duty of the supervisor and is done in cooperation with the principals and teachers of the various schools. This is accomplished by the selection of a committee who represents the various schools, and is responsible for planning the purpose and needs of the program. In choosing the tests to be used, it is necessary that the committee become thoroughly familiar with them. Various questions must be answered, such as: What is the test's general content? Is it easy to administer and score? What is the validity and reliability and are the norms established on a satisfactory standardization population?

After the test has been found suitable within itself, it is then studied to decide if it is practicable for use in local schools. Most important of all, it must serve the specific purposes under consideration.

---

<sup>4</sup>C.C. Ross, Measurement in Today's Schools, p. 186.



Ordinarily the intelligence test to be given is selected before the achievement test, for it is necessary to know to some degree the quality of pupil ability in the group before it is possible to pass judgment on his accomplishments. It is necessary to administer the intelligence test several times during the school career of a child because of differences which might arise in the test, or the unreliability of measurement.

In the St. Petersburg primary schools an intelligence test is administered in the first and in the fifth grades. The Stanford-Binet and Otis intelligence tests have been used in past years, while last year the California Test of Mental Maturity was given. Ideally, intelligence tests should be given every year, and if this is not possible, they should be given in grades one, three, and six. This allows more continuous observance of a child's growth pattern.

The following tests<sup>5</sup> are recommended for use in the primary school:

1. Pintner-General Ability Tests: verbal series.

Pintner Cunningham Primary Test--grade one and  
the first half of grade two.

Pintner-Durost Elementary Test--grades two to  
four.

Pintner Intermediate Test--grades four through six.

---

<sup>5</sup>These tests are published by World Book Company.





2. Otis Quick-Scoring Mental Ability Tests

Alpha test--grades one to four.

Beta test--grades four through six.

3. Otis Self-Administering Tests of Mental

Ability. The Intermediate Examination

--grades four through six.

"Through intelligence tests it is possible to discover the potential ability of every individual, the general level of the group, and the range of ability in the group. The results of an intelligence test have particular significance when there is a marked discrepancy between the intelligence measure and the achievement results" says Hildreth<sup>6</sup> in her Manual for Interpreting.

After the intelligence tests have been selected, the achievement tests must be chosen. Of these, the Stanford Achievement Test, the California Achievement Test, and the Metropolitan Achievement Test are recommended; all three have been used in the St. Petersburg primary schools. The Metropolitan was the one most recently used, and will thus be the test considered here. Brownell<sup>7</sup> has suggested the following criteria for selecting tests:

1. Does the test elicit from the pupils the desired types of mental processes?

---

<sup>6</sup>p. 17.

<sup>7</sup>William A. Brownell, Some Neglected Criteria for Evaluating Classroom Tests. p. 485-486.





2. Does the test enable the teacher to observe and analyze the thought processes which lie back of the pupils' answers?
3. Does the test encourage the development of desirable study habits?
4. Does the test lead to improved instructional practice?
5. Does the test foster wholesome relationships between teacher and pupils?<sup>8</sup>

The Metropolitan Achievement Test was chosen because its Batteries are designed to cover all fields in which achievement should have been made during the year, and because the different Batteries are made up for each grade. Achievement tests cover the advancements made in such subjects as spelling, writing, arithmetic, science, reading, and language usage.

A knowledge of environmental and health factors is of use in interpreting school progress, and should be considered together with the test results. If the test is interesting to the child is an important point, for then the results will tend to be more valid.

Factors such as the purposes of the program, problems to be dealt with, teachers' time and experience, funds available, and the general organization of the school system must be considered in selecting the tests. In planning the testing program it is important that the tests be ordered

---

<sup>8</sup>Ibid., p. 17.



from the publisher early enough so that there will be no delay for this reason. A month, or more is the time usually allowed for this. The tests arrive from the publisher in packages of about twenty-five each, and each package contains the directions for administering, keys and directions for scoring, and a class record and analysis chart. The tests are repackaged at the supervisor's office in Clearwater, Florida, for the various schools in accordance with recent enrollment figures. Great discretion is used in handling the tests at all times, and teachers are instructed to keep the tests, when not in use, in locked compartments.

#### Frequency of Testing

In St. Petersburg, the intelligence test is administered in the early part of January, while the achievement tests are given to grades two, three, four, five, and six, during a period of four days in May, at one sitting per day. By giving these tests in the Spring, the advantages obtained are:

1. Measuring achievement at the point of completion of the work.
2. Showing growth over the years in instances where there is a regular yearly testing program.
3. Providing information for sectioning pupils for the following grade.<sup>9</sup>

---

<sup>9</sup>Gertrude H. Hildreth, Manual for Interpreting. p. 19.





4. Making reports to parents.
5. Helping to determine promotion.<sup>10</sup>

By giving the tests in the Fall also, the additional advantages listed below are gained:

1. Providing an objective record of achievement for every pupil for the entire class at the time when the pupils are new to the teacher.
2. Providing a basis for modifying instructions during the year by revealing those portions of the proposed classwork which need most emphasis.
3. Removing the time pressure which occurs when tests are given at the end of the year.
4. Removing any temptation on the part of the teacher to "coach" for a test.
5. Motivating pupils.<sup>11</sup>

Although the principal is responsible for choosing the day the test is to be given in her own school, the tests are usually administered at the same time in all of the schools due to the guidance of the Central Testing Bureau. Because the testing program is planned well in advance of the day of testing, and the testing dates and hours are worked out by the supervisor, there is much less tendency for the program to conflict with programs and schedules in the various schools. Usually the tests are administered

---

<sup>10</sup>Ibid., p. 19.

<sup>11</sup>Ibid., p. 19.





early in the week so that if there is any reason that more time may be needed, the program may still be completed within one week.

In order to have a good overview of all the girls and boys in Pinellas County elementary schools, the tests are administered to negro, as well as white, pupils. The results are sent by the individual teacher to the principal who then delivers them to the branch office at the Mirror Lake Junior High School. The supervisor receives them there, and is responsible for tabulating the results and for representing them in graphic form.

It is through the cooperation of the supervisor of the testing bureau, the teachers, and the principals of the various schools that the elementary program is planned and carried out. Through a careful selection of tests, trained administrators, and through a systematic and thorough administration of the tests, the Central Testing Bureau is able to maintain a complete record of a child's progress through the primary grades.



## CHAPTER IV

### THE TESTS

The two most recently used tests in the St. Petersburg primary schools, the California Test of Mental Maturity intelligence test and the Metropolitan Achievement Tests will be discussed in this chapter.

The co-authors of the California Test of Mental Maturity are Elizabeth T. Sullivan, Willis W. Clark, and Ernest W. Tiegs, all well known for their numerous articles in psychology and various tests and research studies. The California Test, a diagnostic test of mental maturity is designed for pupils in grades four to eight. Serving as a general measurement of mental maturity, it contains materials suitable for indicating the degree of maturity in relation to the major factors involved in intellectual capacity as well.

The primary purpose of the test is "to make for each person a diagnostic evaluation of those mental abilities which are related to, or determine, his success in various types of school activity, in order that the teacher may utilize this information directly in aiding pupils who have learning difficulties."<sup>1</sup> Based upon philosophy, research, the work of the authors and of other prominent leaders in the field of mental measurement, the test shows by means of

---

<sup>1</sup>Manual of Directions, California Test of Mental Maturity--Elementary Series. p. 1.





a profile the extent to which a pupil uses his abilities. This makes it possible to see immediately the probable sources of difficulty or success of the pupil, and to provide the guidance necessary. Results are used to form three mental ages and intelligence quotients--"a non-language M.A. and I.Q., a language M.A. and I.Q., and the M.A. and I.Q. characteristic of the familiar intelligence test."<sup>2</sup>

Each battery of the test is preceded by tests of visual acuity, auditory acuity, and motor coordination. Through these tests pupils with defects serious enough to effect scores in the following test are detected. In building the test batteries recognized mental processes or functions are taken into account and present knowledge of the nature of mental growth and development are utilized.

The tests, in general, sample the maturity of memory, of apperceptive processes, of spatial relationships, and of logical and mathematical aspects of reasoning. Some of the tests are verbal, others non-language; thus making it possible to obtain a separate evaluation of these two aspects of mental maturity.<sup>3</sup>

The Pre-Tests listed in the California Mental Maturity Test are as follows: visual acuity, auditory acuity, motor coordination; the tests of mental maturity are: memory-immediate recall, memory-delayed recall, spatial relation-

---

<sup>2</sup>Ibid., p. 2.

<sup>3</sup>Ibid., p. 2.





ships, logical reasoning, numerical reasoning, and vocabulary. The test is given in two periods of about fifty minutes each. It is a power rather than a speed test and only one test calls for a definite time limit--the eighth, which requires that the child draw a continuous line on a drawing without crossing or going outside the lines.

The test has a high degree of reliability as is shown in the table below:<sup>4</sup>

	---Grades---		
	4	5	6
Total mental factors.....	.956	.950	.958
Language factors.....	.930	.946	.944
Non-Language factors.....	.925	.922	.932
Test A. Memory.....	.906	.933	.926
Test B. Spatial Relationships....	.873	.898	.900
Test C. Logical Reasoning.....	.874	.876	.887
Test D. Numerical Reasoning.....	.894	.889	.897
Test E. Vocabulary.....	.912	.922	.901
No. of Pupils.....	144	270	308
S.D. (M.A. in months).....	12.0	13.0	13.5

"The probable error of estimate for an individual pupil score varies from two months of mental age when the reliabilities for single grade are .95 and above to four months when the reliabilities are .87 to .89."<sup>5</sup>

The California Test of Mental Maturity compares favorably with the Stanford-Binet with a correlation of .88. In regards to the validity of the test, the authors state that:

The practice of dealing only with mental ages and intelligence quotients obscures and<sup>6</sup>

---

<sup>4</sup>Ibid., p. 2.

<sup>5</sup>Ibid., p. 2.

<sup>6</sup>Ibid., p. 2.



ignores the separate important factors which constitute mentality; and it is in terms of these factors that the abilities of individuals should be diagnosed. This battery attempts such a diagnosis on the basis of present knowledge; further evidence of validity must await further knowledge of the nature of mental maturity.<sup>7</sup>

The authors define the Intelligence Quotient as a fairly constant ratio and state that it indicates the degree of ability that the individual has in comparison with others of his age. I.Q.'s for this test are ordinarily interpreted as follows:<sup>8</sup>

I.Q.	Descriptive Classification	% of Typical pop.
130 and above	Very superior	3%
115 - 129	Superior	12%
100 - 114	High average	35%
85 - 99	Low average	35%
70 - 84	Inferior	12%
Below 70	Very inferior	3%

Buros<sup>9</sup> states that Garret, in criticizing the California Test of Mental Maturity, says that the authors offer no evidence to prove that the tests are constructed to fulfill conditions of I.Q. constancy. However, in favor of the test, he states that the information is complete, the data adequate, that the pictorial tests are clear, and that it is a well-made test. The test is valuable to the school as an element in the counseling and guidance program, and as basic information in

---

<sup>7</sup>Ibid., p. 2.

<sup>8</sup>Ibid., p. 6.

<sup>9</sup>Buros, Mental Measurements Yearbook.





fitting instruction to the needs of the pupil<sup>10</sup>.

### The Metropolitan Achievement Tests

The Metropolitan Achievement Tests series was first begun in the early 1930's with the construction of a series of tests intended primarily for New York City. Almost immediately after the New York City edition was published, a new series for national distribution was introduced. In 1936 the tests were equated to the Stanford Achievement Tests: Forms V, W, X, Y, and Z; and since that time they have been used in all parts of the country<sup>11</sup>.

Since 1936, however, the school curricula has been undergoing a constant change. A need arose for tests containing new material, and the Metropolitan series was revised to meet this need. "A national program was carried out as a cooperative venture on the part of the participating communities, the authors, and the publisher of the tests."<sup>12</sup> In 1946, over 500,000 tests were given in a national testing program carried out in every state in the country. The result of this national standardization program was the introduction of the forms now in use--R, S, T, U, and V.

The Metropolitan Achievement Tests series consists of

---

<sup>10</sup>Manual of Directions, op. cit., p. 7.

<sup>11</sup>Gertrude H. Hildreth, Manual for Interpreting., p. 92.

<sup>12</sup>Ibid., p. 8.





five batteries covering the essential skill subjects and content areas taught in grades one to nine. There is also available the Metropolitan Readiness Tests which help determine the readiness of pupils to learn first grade work. Each battery consists of a group of different tests which were standardized at the same time, on the same pupil population<sup>13</sup>. The following chart shows the content of the batteries used in grades one through six<sup>14</sup>:

Test Unit	Grade(s)	Subtests
Primary I Battery	1	word pictures, meaning, recognition, and numbers
Primary II Battery	2	Reading, word meaning, arith. fund., arith. prob., spelling
Elementary Battery	3 & 4	Reading, vocab., arith. fund., arith. prob., language usage, spelling
Intermediate Battery Complete	5 & 6	Read., vocab., arith. fund. and prob., Eng. Lit., geog., hist., civics, science, spelling
Intermediate Battery Partial		Read., vocab., arith. fund. and prob., English, spelling
Elementary Reading Test	3 & 4	Reading and vocabulary
Int. Reading Test	5 & 6	Reading and vocabulary
Elementary Arith. Test	3 & 4	Arith. fund. and arith. prob.
Int. Arith. Test	5 & 6	Arith. fund. and arith. prob.

Particular attention has been given to the drawings in the Primary Batteries. The forms contain drawings for every item, with emphasis on clarity and freedom from ambiguity and every effort has been taken to make the drawings interesting to the children<sup>15</sup>.

---

<sup>13</sup>Ibid., p. 5.

<sup>14</sup>Ibid., p. 5.

<sup>15</sup>Ibid., p. 10.



Tests in the Metropolitan Achievement series are not valid in general, but they are valid for specific purposes.

"The authors consider this series of tests to be valid as a measure of typical content of instruction in this country. As such it provides a local community using the tests with a stable reference point in terms of which to evaluate its own achievement."<sup>16</sup>

It is possible to justify all material used in the tests in terms of representative courses of study, popular textbooks, and opinions of experts in the various fields<sup>17</sup>.

The Metropolitan Achievement Tests have a high reliability coefficient, as is shown in the table on the following page (p. 39). Because of its high degree of reliability and validity, and because of the ease of administering and scoring the tests, the Metropolitan Achievement series is considered one of the best achievement tests for the primary grades; and when the results of both the achievement tests and the intelligence tests are applied correctly in helping and guiding the child through the primary grades, the results are certain to be satisfactory.

---

<sup>16</sup> Gertrude H. Hildreth, Manual for Interpreting., p. 8.

<sup>17</sup> Ibid., p. 8.



Corrected Split-Half Reliability Coefficients, by the Spearman-Brown formula, for the Metropolitan Achievement Tests:  
Form R. (based on three communities of similar character)<sup>18</sup>

Battery	Grade	N	Test	R.C.
Primary I	1	402	Word Pictures	.953
			Word Recognition	.892
			Word Meaning	.878
			Numbers	.951
Primary II	2	363	Reading	.951
			Word Meaning	.947
			Arith. Fund.	.970
			Arith. Prob.	.852
			Spelling	.936
Elementary	3	374	Reading	.959
			Vocabulary	.927
			Arith. Fund.	.946
			Arith. Prob.	.871
			Language Usage	.924
			Spelling	.934
Intermediate	5	350	Reading	.954
			Vocabulary	.926
			Arith. Fund.	.914
			Arith. Prob.	.879
			English	.904
			Literature	.859
			History & Civics	.789
			Geography	.806
			Science	.821
			Spelling	.933

---

<sup>18</sup>Ibid., p. 9.





## CHAPTER V

### ADMINISTERING AND SCORING THE TESTS

#### Administering

The proper administration of the tests is of utmost importance for the reason that for comparative purposes the tests are of value only if their directions are carried out with exactness. The first step to consider therefore is who should administer the tests?

Only competent persons should be allowed to give the tests. In the case of some tests, such as the Stanford-Binet, special training is required of the administrators. The reason for this is that in order for the results to be most valid it is necessary that the administration and scoring of the tests be as unvarying as possible. In the primary schools of St. Petersburg it is the classroom teacher who is responsible for giving the tests to the children. In preparation for this, it is the requirement of the schools that an elementary teacher should have at least two years of college, be certified, and should have had a course in testing and/or statistics.

It is not well to take the competency of the examiner for granted, however. If possible, the tests to be used should be demonstrated so that any questions arising concerning their administration may be suitably answered. In St. Petersburg, the principal calls a conference of the



teachers. Here, the purposes of the tests are explained and the teacher is supplied with the necessary materials. Hildreth has stated in her "Manual for Interpreting" an ideal situation as follows:

....The ideal situation is to have a series of such conferences: the first, before the administration of the tests, for the purpose of orienting the personnel in test administration; the second, after the administration of the tests and before the scoring, for the purpose of setting up the best scoring procedure; and the third, after the scoring is completed, for the purpose of discussing the preparation and interpretation of test results. This arrangement is not possible under all conditions and accordingly one or two meetings can be made to serve the needed purposes.<sup>1</sup>

The procedure of administering group intelligence tests and achievement tests is not difficult, yet certain measures must be attended to in order to insure satisfactory results. The best place for the tests to be given is in the familiar surroundings of the pupils' own classroom. The best time is during a regular class period so that the testing time will not run over into the lunch or play hour. Also, it is not desirable to give the tests before an important event such as a holiday or class program, as full attention will not be devoted to the tests. In order to insure freedom from distractions and interruptions a sign should be posted on the door informing of the tests going on inside. Just

---

<sup>1</sup>Gertrude Hildreth, Manual for Interpreting, p. 21.





before test-time pupils are instructed to remove everything from their desk tops other than the materials which they will use for the test.

In the actual administration of the test there are three important requirements which the teacher must meet.<sup>2</sup> The first is the ability to read silently. This is necessary so that the directions printed in the manual accompanying the tests may be mastered. Good oral reading is necessary also, so that the directions may be read to the pupils rather than recited from memory. By reciting from memory the teacher runs the risk of omitting some important words or phrases. Directions should be read with proper emphasis in a low, clear voice so that the possibility of arousing anxiety or excitement is avoided.

Secondly, the teacher must be able to keep time accurately. A stop watch is used, and the time must be recorded accurately as to the hour, minute, and second which the test was begun and stopped. The time should be recorded immediately and never trusted to the memory.

The third requirement is that the teacher must be able to follow directions accurately. The test manual should be followed to the letter, for if any changes are made the test will no longer be considered standardized.

In administering the test the teacher should keep

---

<sup>2</sup>C.C. Ross, Measurement in Today's Schools. pp. 204-206.





in mind that the test is actually a standardized situation in which to observe behavior. She should write down any information which might help in interpreting the results, should be constantly on the alert to see that pupils do not help one another, and should be ready to answer questions rapidly so that the pupil will lose no more time than is necessary. "In case of doubt, the examiner should be on the side of saying nothing."<sup>3</sup> This will eliminate all chance that time will be wasted or improper information given out.

### Scoring

After the tests have been given, it is desirable that they be scored as quickly as possible and with strictest accuracy, for no matter how well the tests have been administered, the results are worthless if they are inaccurately scored. Studies have been made by Morrison, Pintner, Madsen, and Dearborn and Smith<sup>4</sup> into the errors made on tests. Their investigations resulted in the revealing of two types of errors in scoring; constant errors and variable errors. Constant errors are the result of such mistakes as misunderstanding the directions for scoring, or by counting omissions as errors. Variable errors result from carelessness, errors in entering scores on record sheets, and errors in addition.

---

<sup>3</sup>Ibid., p. 206.

<sup>4</sup>Ibid., p. 207.



The only way to guard against these errors is through prevention by teaching the administrators the proper techniques to be employed, and by double-checking the scores.

Scoring should be as mechanical as possible. In St. Petersburg Primary Schools the teacher is responsible for scoring her class tests and for recording them, in order of highest to lowest score received, on a chart on the front of the package in which the tests were received.

An impersonal attitude should be maintained throughout the scoring of the tests so that the papers will be scored more accurately. When one teacher works alone in scoring the tests for her class she should follow a pattern of scoring one or two sub-tests for all the pupils in the class at a time. This will increase the speed of scoring, and she will become familiar with the procedure as she goes along.

Another method of scoring is that of working in a group. In the group method, a number of teachers work together, and each teacher is responsible for one or two tests. This system places the scoring on a clerical basis; an impersonal attitude is maintained; and the process speeds up the scoring. This method is used in the larger elementary schools in Pinellas County.

After the tests have been scored, it is necessary that they be rechecked. It is not necessary to check all papers but rather to take about five papers at random, check





them, and if no errors are found, consider the rest satisfactory. If more than two errors are found on any one test, another group of papers should be rechecked for that particular test. If more than two errors occur on these five, all the tests should be rescored. This procedure should be followed for all sections of the test. Figures should be rechecked every time they are transferred from one paper to another, and again, accuracy is the prime factor.

After the scores have been recorded the results are recorded and may then be interpreted. The values derived from the proper use of the results are discussed in the following chapter. Therefore, a few of the precautions which should be observed in the use of test results will be mentioned here.

Because no test is perfectly reliable, the results should not be the only means of rating the pupil. In order to administer remedial action correctly, it is advisable that the pupil be retested with another test, and the results compared. In comparing the pupils with national norms the individual, his environment, and background should be considered. These considerations should also be noted in comparing classes and schools.

Sometimes the teacher is rated by the test results. This is, of course, not a good policy as teachers will often go to extremes in using unethical methods to insure higher scores if they think their teaching reputation is at stake.





In St. Petersburg primary schools the results of the testing program have no effect upon the teachers' methods of instructing her class other than to be used as a basis of comparison in order to discover better methods of teaching.

The testing program is of little use if it is not followed up with remedial practices. These should be begun as soon as possible after the tests are administered and scored. In St. Petersburg each class in the primary grade uses some kind of follow-up procedure; usually classes in reading are divided into slow, normal, and rapid readers. The "Alice and Jerry" series of readers, which is adapted for this type of instruction, are used. Arithmetic classes are also divided into small groups according to the pupils' capabilities. General principles concerning individualized help are:<sup>5</sup>

1. The use of the class distribution based on test results to locate the slow learners.
2. The analysis of errors made by the entire class to locate the most common error.
3. The analysis of individual papers as a basis for individualized instruction.
4. The organization of small-group work for pupils who are at similar levels of achievement and have somewhat the same difficulties.
5. The use of concrete activities and realistic learning experiences to enrich meanings of

---

<sup>5</sup>Gertrude H. Hildreth, Manual for Interpreting. p. 91.



6. The use of the teacher's own observations along with the test data to formulate plans for individualized help.<sup>6</sup>

Both proper administration and proper scoring must be employed to receive from the tests the results which are expected. If teachers are properly trained, and if they use the care demanded in following out the test manual in administration and scoring procedures, these results will most probably be obtained.

---

<sup>6</sup>Ibid., p. 91.



## CHAPTER VI

### THE VALUE OF TESTING IN THE PRIMARY SCHOOL

After the tests have been carefully administered and scored the results are then ready to be put to use. The results are recorded, tabulated, and are ready to serve their purpose by being presented in a meaningful way which will arouse interest on the part of those most concerned--the pupils, the teachers, the administrative officers, and the parents.

In achievement tests, the pupils should be informed as to their performance on the tests, although they should not be told their scores. By going over the errors with the pupils they will see their mistakes, and steps may be taken to prevent their re-occurrence. In the case of intelligence tests, however, the results should not be made known to the child. It is very hard for a child to understand the reason he makes a low score on such a test, and he would tend to feel inferior to those children with higher scores. If the result of a pupil's intelligence test is a high score, there is the danger that the child will become overly conceited if his score is made known. In any case, utmost discretion must be employed when making the results known. The chief value a child will derive from knowledge of achievement test results is that he may strive to do better and to improve himself so that he will make a higher





score the next time.

By learning the Intelligence Quotient of the child from the Intelligence tests a teacher is able to form an idea of what his rate of achievement should be. This is an undesirable factor if the teacher misconstrues the purpose of the test, and tends to expect too much of a child of a high I.Q., or drives the child of a low I.Q. beyond his capacity. Some children, the teacher will find, are slow learners. Individualized instruction based on a study of the child's needs will help him make slow, but continuous, progress in terms of his needs and capacities as a learner. Other children, although they have normal learning capacities, will be deficient in certain fields but can also be helped through individual instruction.

The teachers use the test results to learn more about the individual pupils they teach. Through standardized tests teachers realize that pupils may be outstanding in some subjects, but may be very weak in others. Uses to which teachers might put test results are listed by Hildreth<sup>1</sup> as follows:

1. To determine the achievement level of each pupil in each subject with relation to age and ability, with equal emphasis on the gifted as well as on the slow or special-problem cases.

---

<sup>1</sup>Manual for Interpreting, p. 46.



2. To compare present achievement with past achievement in order to determine the rate of progress.
3. To determine the average achievement level of a class at the beginning or end of the year.
4. To obtain a picture of the nature and range of individual differences in the group.
5. To group pupils within the class for instruction.
6. To diagnose an individual pupil's difficulties in learning.
7. To call attention to subject areas where more detailed testing is required to check on the effectiveness of teaching.
8. To provide a basis for counseling with parents regarding a pupil's achievement.<sup>2</sup>

Each teacher carefully records the test results on the back of each pupil's permanent record folder. The name of the test, the date the test was given, and the grade in which the pupil was when the test was given are all noted. Some of the tests contain an individual graph. These graphs help the teachers in that they instantly show strengths and weaknesses of the child. The graphs are also placed in the permanent record folder. These data are valuable as they show a picture of the child's growth and development. Through the achievement test date, the intelligence test results, school records, health records, and environmental conditions, the teacher is able to develop a better understand-

---

<sup>2</sup>Ibid., p. 46.





ing of the particular type of instruction needed for each individual in her class.

For a study of the class as a whole and for further study of the individual, the teacher makes up a Class Record from which she composes a Class Analysis Chart. This chart shows the median of the grade and how the results of the achievement tests vary in the different batteries. It is also used to show the distribution of chronological ages, mental ages, and Intelligence Quotients. The completed Analysis Chart is valuable in that it aids the teacher in:

- a. Studying the achievement of the class in the light of intelligence by comparing the median achievement age on each test with the median mental age.
- b. Identifying pupils whose achievement is not in line with intelligence by comparing the achievement age in each subject with mental age.
- c. Identifying subject strengths and weaknesses by comparing the median achievement in each subject area with the local or national norm.
- d. Ascertaining the range of abilities in each of the subject areas in comparison with the range of intelligence by comparing achievement ages with mental ages. . . .
- e. Identifying individual pupils who are in need of special attention as a basis for remedial instruction by locating those pupils who are significantly above or below the median of the class.
- f. Planning instruction in the light of education-<sup>3</sup>

---

<sup>3</sup>Ibid., pp. 55-58.





al needs and abilities of the class by modifying the instructional program for some pupils or by grouping pupils for instructional purposes.

- g. Motivating pupils toward self-development by identifying their performance in relation to the performance of the class. The use of identification numbers rather than names makes this type of comparison strictly anonymous.<sup>4</sup>

Another value of testing early in the child's school career is that many children have minor physical defects which may not have been discovered, yet they slow down the natural development of the child. These children with physical defects do not work up to the mental ability of which they are capable, and the tests will point to the fact that something is wrong. Poor hearing and eyesight are examples of such a disability.

Testing will help the principal of the school learn more about the weaknesses of the school structure so that administrative changes may be made if necessary. For example, if achievement is found to be especially low in a particular subject, it will be possible to study teaching practices, class groupings, and textbooks to find where the change must be made. If the performance of one particular class is outstanding, it may be studied as a guide to improve the teaching of other classes.

The results of the tests, when evaluated by the super-

---

<sup>4</sup>Ibid., pp. 55-58.



visor of the Central Testing bureau, are valuable to determine the needs for improved teaching methods. The supervisor of the testing program for Pinellas County prepares each year a bulletin to be distributed to the teachers of the various elementary schools. Entitled "A Summary of the Testing Program in the White Elementary Schools During the School Year \_\_\_\_ to \_\_\_\_", it shows subject medians for the school system and the strengths and weaknesses among the pupils. From graphs in the bulletin the teacher can see at a glance the percentage of pupils expected to complete school, as far ahead as college preparatory courses, and also the percentage of pupils who will need a vocational type of education. It is possible, too, for the teacher to check to see if her class is capable of doing work equal to the county median, or whether they should do better than average.

Parents profit by the administration of the tests because much of the maladjustment problems of pupils in later years is the result of the fact that their parents expect too much of them. If a parent would find out the level of the child's intelligence before encouraging him to develop levels of accomplishment beyond his capabilities, many of these maladjustment problems would be averted. If the child's Intelligence Quotient is lower than that which is necessary for success in school, he can be directed into the proper channels so that he can be properly guided. Although the



St. Petersburg schools are not so equipped, some schools have classes for 'exceptional children' where the pupils are taught to develop within their abilities.

Because intelligence and achievement testing programs have proved their worth in schools which have experimented with them, more and more schools throughout the country are following suit. The better adjusted children, and the more understanding teachers and parents resulting from the testing programs best show the value of such a program in the elementary schools.





## CHAPTER VII

### SUMMARY AND CONCLUSIONS

After making a careful investigation into the testing program of the St. Petersburg, Florida, primary schools, with attention directed toward the use of the intelligence and achievement tests in the first six grades, the writer has shown that the schools are definitely 'test-conscious', and that all the school personnel, from the supervisor<sup>1</sup> of the testing program to the principals<sup>2</sup>, to the teachers<sup>3</sup> and to the parents<sup>4</sup> are interested in the pupil as an individual and in his needs.

The testing program is county-wide, and is centralized around the Central Testing Bureau located in Clearwater, Florida. The study has shown that such a Bureau is one of value for any school which offers a testing program in that:

1. It is headed by one person who is in contact with school principals and teachers. She obtains their aid in selecting tests which aid them most in their teaching, guidance, and supervision, through a

---

<sup>1</sup>Supra., p. 23.

<sup>2</sup>Supra., p. 52.

<sup>3</sup>Supra., p. 49.

<sup>4</sup>Supra., p. 53.



series of conferences.

2. Tests are ordered from, and distributed by the Central Testing Bureau. This insures a uniformity throughout the county.
3. Test results are returned to the bureau where they are tabulated and the results printed in a bulletin which is distributed to the teachers. The bulletin serves as a guide for the teacher in comparing her class average with other schools, and thereby remedying any weaknesses.
4. Teachers are trained by Bureau personnel to administer and score tests correctly.
5. The Bureau sets the days for the tests to be administered. This does away with a great deal of confusion.

The study has shown there is a definite need for good tests in the school. All children show learning differences in school as well as in all other traits. Pupils of similar ages do not progress at the same rate, even with the best of teaching. Some pupils have physical or mental handicaps; dull pupils, although deficient in comparison with school standards may be accomplishing as much as they are able; others, with the ability to progress at a normal rate, fail to do so. The administration of intelligence and achievement tests brings forth these deficiencies in their early stages,



and with the aid of the teacher, proper remedial measures are taken so that the child is saved from becoming maladjusted in later years. The need and value of a testing program in the first six grades is thus quite evident.

Intelligence tests are administered in the first and the fifth grades in the St. Petersburg primary schools. The investigation has shown that if they could be given more often--preferably every year, or at least in grades one, three, and six--a more reliable measurement of the child would be obtained. More frequent administration of the tests would make it possible to offer a larger selection of tests, so that through comparison of the different test results, the child could be compared with a greater cross-section of school children his own age.

Achievement tests are administered once yearly, at the end of the school term. When tests are given both fall and spring<sup>5</sup>, a better picture of the child is formed by the teacher by checking his achievement during the period from one testing period to another.

Tests are given under controlled conditions by the teachers who have been trained in administration and scoring procedure<sup>6</sup>. In a clear low voice she gives the directions, times the tests with precision, and gives aid to the pupils

---

<sup>5</sup>Supra., pp. 29-30.

<sup>6</sup>Supra., Chapter V.





when necessary. By following directions exactly, more valid results are obtained, and the tabulated results of all the St. Petersburg elementary schools are thus more reliable.

Scoring<sup>7</sup> is done by the teacher, or by a group of teachers, depending upon the size of the school. Results are checked and rechecked to insure correctness.

After scoring is completed, the results are entered on the pupils' Cumulative Record. This record goes with the pupil as he goes through school. Future test scores will be recorded so that other teachers will be better prepared to understand the child, and thus offer him proper guidance.

If proper intelligence and achievement tests had not been administered in the primary school, and remedial treatment begun when necessary, many a child would not progress as far in school as do those children who are guided from the first grade on.

---

<sup>7</sup>Supra., pp. 43-47.



## BIBLIOGRAPHY



- Bainum, Mary I., A Summary of the Testing Program in the White Schools During the School Year 1947-1948. Clearwater, Florida. February, 1949.
- Baker, Harry J., Introduction to Exceptional Children. New York: MacMillan Company, 1948.
- Barr, A.S., Burton and Brueckner, L.J., Supervision. New York: D. Appleton-Century Company, 1938.
- Boynton, Paul L., Intelligence, Its Manifestation and Measurement. New York: D. Appleton-Century Company, 1933.
- Broom, M.E., Educational Measurements in the Elementary School. New York: McGraw-Hill Book Company, 1939.
- Brownell, William A., Some Neglected Criteria for Evaluating Classroom Tests. National Elementary Principal, July, 1937.
- Brubacher, John S., A History of the Problems of Education. New York: McGraw-Hill Book Company, 1947.
- Buros, Oscar K., Mental Measurements Yearbook of the School of Education. New York: Rutgers University Press, 1937.
- Caldwell, Otis W., and Courtis, Stuart A., Then and Now in Education. New York: World Book Company, 1942.
- Crow, Lester D., and Alice, Mental Hygiene in School and Home Life. New York: McGraw-Hill Book Company, 1942.
- Douglas, Harl R., The High School Curriculum. New York: The Ronald Press Company, 1947.
- Freeman, Frank N., Mental Tests. New York: McGraw-Hill Book Company, 1939.
- Fuguitt, G.V., and Lynch, George M., A Study of Pupil Achievement in Pinellas County Public Schools. Clearwater, Florida, 1932.
- Garrett, Henry E., Great Experiments in Psychology. New York: D. Appleton-Century Company, Inc., 1941.
- Greene, Edward B., Measurements of Human Behavior. New York: The Odessey Press, 1941.
- Greene, Harry A., and Jorgenson, A.N., The Use and Interpretation of Elementary School Tests. New York: Longmans, Green and Company, 1935.





Hildreth, Gertrude, A Bibliography of Mental Tests and Rating Scales. New York: Psychological Corporation, 1933.

\_\_\_\_\_, Manual for Interpreting. New York: The World Book Company, 1948.

Jones, Arthur J., Principles of Guidance. New York: McGraw-Hill Book Company, Inc., 1945.

Kandel, I.L., Examinations and Their Substitutes in the United States. New York: The Carnegie Foundation for The Advancement of Teaching.

Kelly, T.L., Interpretation of Educational Measurements. New York: World Book Company, 1927.

Kuhlman, F., Tests of Mental Development. Minneapolis, Minn., Educational Publications, 1939.

Lee, J. Murray, The Child and His Curriculum. New York: D. Appleton-Century Company, Inc., 1940.

Madsen, I.N., Participation in Testing Programs by the Classroom Teacher. Educational Administration and Supervision, February, 1929.

McCall, William A., Measurement. New York: MacMillan Company, 1939.

Monroe, Walter S., Educational Tests and Measurements. Cambridge, Mass.: The Riverside Press, 1917.

Munn, Norman L., Psychology. New York: Houghton Mifflin Company, 1946.

Nelson, M.J., Tests and Measurements in Elementary Education. New York: The Cordon Company, Inc., 1939.

Orleans, Jacob S., Measurement in Education. New York: Thomas Nelson and Sons, 1937.

Pintner, Rudolph, Intelligence Testing. New York: Henry Holt and Company, 1927.

Remmers, H.H., and Gage, N.L., Educational Measurement and Evaluation. New York: Harper and Brothers, 1943.

Ross, C.C., Measurement in Today's Schools. New York: Prentice-Hall, Inc., 1941.



- Spence, Ralph B., A Comprehensive Testing Program for Elementary Schools. Teachers College Record, January, 1933.
- Stoddard, George D., The Meaning of Intelligence. New York: MacMillan Company, 1943.
- Sullivan, Elizabeth T., Clark, Willis W., and Tiegs, Ernest W., Manual of Directions--California Test of Mental Maturity--Elementary Series. Los Angeles, California: California Test Bureau, 1946.
- Terman, Lewis B., The Measurement of Intelligence. New York: Houghton Mifflin Company, 1926.
- Terman, L.M., and Merrill, M.A., Measuring Intelligence. Boston, Mass., Houghton Mifflin Company, 1926.
- Thoburn, Ralph, H., The Measurement and Improvement of Pupil Adjustment. Washington, D.C.: National Education Association, sixteenth yearbook of the Department of Elementary School Principals.
- Thorndike, E.L., The Measurement of Intelligence. New York: Teachers College, Columbia University, 1927.
- Thurston, L.L., Primary Mental Abilities. Chicago, Illinois: University of Chicago Press, 1938.
- Tiegs, Ernest W., and Clark, Willis W., Progressive Achievement Tests--Elementary Battery. Los Angeles Test Bureau, 1937.
- Traxler, Arthur E., Techniques of Guidance. New York: Harper and Brothers, 1945.
- Ulich, Robert, History of Educational Thought. New York: The American Book Company, 1945.
- Webb, L.W., and Shotwell, Anna M., Testing in the Elementary School. New York: Farrar and Rinehart, 1939.
- Whipple, G.M., Manual of Mental and Physical Traits. Baltimore, Maryland: Warwick and York, 1910.
- Woody, Clifford and Sangren, Paul V., Administration of the Testing Program. New York: World Book Company, 1932.
- Wrightstone, J. Wayne, Appraisal of Newer Elementary School Practices. New York: Teachers College, 1938.













